BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 7: e111048
doi: 10.3897/biss.7.111048

OPEN ACCESS

Conference Abstract

A Case Study in Sharing Marine eDNA Metabarcoding Data to OBIS

Katherine Silliman[‡,§], Sean Anderson[|], Rachael Storo[‡,§], Luke Thompson[‡,§]

‡ NOAA Atlantic Oceanographic and Meteorological Laboratory, 4301 Rickenbacker Cswy, Miami, FL 33149, United States of America
§ Northern Gulf Institute, Mississippi State University, Starkville, MS, United States of America
| Department of Biological Sciences, University of New Hampshire, 38 Academic Way, Durham, NH 03824, United States of America

Corresponding author: Luke Thompson (luke.thompson@noaa.gov)

Received: 11 Aug 2023 | Published: 15 Aug 2023

Citation: Silliman K, Anderson S, Storo R, Thompson L (2023) A Case Study in Sharing Marine eDNA Metabarcoding Data to OBIS. Biodiversity Information Science and Standards 7: e111048. https://doi.org/10.3897/biss.7.111048

Abstract

Metabarcoding of DNA collected from an environmental sample (eDNA) is increasingly employed in marine biodiversity surveys, with the ability to target taxa from microbes to plankton to large vertebrates depending on the molecular markers used. These techniques are often the only viable method to detect certain taxonomic groups, and therefore provide observations that are currently under-represented on existing biodiversity data platforms, such as the Ocean Biodiversity Information System (OBIS) and the Global Biodiversity Information Facility (GBIF). Some of the reasons for this disconnect include the unique data structures inherent to eDNA datasets, the complexities of combining marine observation data and environmental data (De Pooter et al. 2017), and the minimal availability of documented examples.

Here we present a detailed case study on the preparation of marine metabarcoding survey data for publication to OBIS. This data comes from the 2021 Gulf of Mexico Ecosystems and Carbon Cycle (GOMECC) cruise led by the National Oceanic and Atmospheric Administration (NOAA), employing 17 coastal-offshore transects across the Gulf of Mexico and the Atlantic Ocean. Metabarcoding libraries targeted bacteria and archaea with the 16S rRNA marker and eukaryotes with the 18S rRNA marker. Amplicon sequence variants (ASVs) were inferred for each marker, then taxonomy was assigned to these ASVs using

© Silliman K et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

the open-sourced reference databases PR^2 5.0.1 and SILVA 138.1. OBIS requires that taxonomic assignments are converted to the World Register of Marine Species (WoRMS) nomenclature, which can be particularly challenging for marine bacteria and archaea as these taxa are underrepresented on WoRMS. Three tables—the per-sample ASV observation counts, assigned taxonomy of the ASV sequences, and sample collection data—were then converted to the DNA derived extension for Darwin Core (Abarenkov et al. 2023) using a combination of new and OBIS-provided Python scripts (LaScala-Gruenewald et al. 2021). This workflow is available at the NOAA Omics Data Management Guide*[1], which will also host links to NOAA Omics datasets.

Some of the key challenges to navigate when preparing metabarcoding data for OBIS include:

1. developing a sample collection data template,
2. converting taxonomic assignments to WoRMS nomenclature while preserving the original assignments, and
3. recording the complex methodological and bioinformatic processes involved in data generation in order to be reproducible.

We describe our workflow for tackling these challenges, with the aim of fostering discussion on best practices for publishing marine eDNA data to biodiversity data platforms. This work is part of a larger effort across NOAA 'Omics to develop a comprehensive bioinformatics platform and data management framework for marine eDNA and microbiome data.

Keywords

DNA barcodes, marine biodiversity monitoring, ASV sequences, Darwin Core

Presenting author

Katherine Silliman

Presented at

TDWG 2023

Acknowledgements

The authors would like to thank Stephen Formel and Abigail Benson for their guidance o working with OBIS.

Funding program

This work was supported by award NA21OAR4320190 to the Northern Gulf Institute, by NOAA Ocean Exploration via the Northern Gulf Institute, and by the NOAA Ocean Acidification Program (project number 21392), each from NOAA's Office of Oceanic and Atmospheric Research, U.S. Department of Commerce.

Hosting institution

NOAA Atlantic Oceanographic and Meteorological Laboratory and the Northern Gulf Institute

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Abarenkov K, Andersson AF, Bissett A, Finstad AG, Fossøy F, Grosjean M, Hope M, Jeppesen TS, Kõljalg U, Lundin D, Nilsson RN, Prager M, Schigel D, Suominen S, Svenningsen C, Frøslev TG (2023) Publishing DNA-derived data through biodiversity data platforms. Copenhagen: GBIF Secretariat. v1.3 https://doi.org/10.35035/doc-vf1a-nr22.
- De Pooter D, Appeltans W, Bailly N, Bristol S, Deneudt K, Eliezer M, Fujioka E, Giorgetti A, Goldstein P, Lewis M, Lipizer M, Mackay K, Marin M, Moncoiffé G, Nikolopoulou S, Provoost P, Rauch S, Roubicek A, Torres C, van de Putte A, Vandepitte L, Vanhoorne B, Vinci M, Wambiji N, Watts D, Klein Salas E, Hernandez F (2017) Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences. Biodiversity Data Journal 5 https://doi.org/10.3897/bdj.5.e10989
- LaScala-Gruenewald D, Pitz K, Chavez F (2021) https://github.com/iobis/dataset-edna

Endnotes

*1 https://github.com/aomlomics/omics-data-management